КАЗАНСКИЙ ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
АВИАЦИОННЫЙ ИНСТИТУТ им. А. Н. ТУПОЛЕВА

АВТОМАТИЗИРОВАННЫЕ ОБУЧАЮЩИЕ СИСТЕМЫ

Межвузовский сборник

КАЗАНЬ
1979

В сборнике освещаются результаты исследований по комплексной разработке автоматизированных обучающих систем, вопросы их психолого-педагогического обоснования и эффективности применения в учебном процессе, а также вопросы, связанные с информационным, математическим и техническим обеспечением других современных технических средств обучения. В сборнике нашли отражение результаты исследовательской работы "Методы применения АОС" и "Комплексная разработка систем автоматизированного диалога и коллективного обучения", проводимых по приказу Минвуза РСФСР.

УДК 31:[15:371]

ПРОБЛЕМА ИЗМЕРЕНИЙ В АДАПТИВНОМ ОБУЧЕНИИ С ПОМОЩЬЮ АВТОМАТИЗИРОВАННЫХ ОБУЧАЮЩИХ СИСТЕМ

С.Ф.Занько, И.М.Юсупов

Применение ЭВМ для целей обучения непременно предполагает измерение тех или иных параметров процесса обучения и характеристик обучаемых. С этой точки зрения современные обучающие системы можно рассматривать как измерительные устройства, предоставляющие большие возможности сбора и переработки информации о процессе обучения и самих обучаемых. Однако практическое использование этих устройств наталкивается на значительные трудности, ибо "вопросов об учебном процессе сущест-

вует такое великое множество и машине можно задать столько вопросов и получить от нее столько ответов, что учителя, администрация и исследователи рискуют быть просто захлестнутыми столь большим объемом статистического материала, что вряд ли смогут с ним справиться" [13, с.173].

Имеется два очевидных выхода из этой ситуации: во-первых, сократить число регистрируемых данных; во-вторых, привести процедуру обработки данных к форме, пригодной для практического использования. Первый из них связан с ответом на вопрос „что измерять?" Второй – с ответом на более трудный вопрос "как измерять?"

Применительно к АОС вопрос о том, "что измерять", модифицируется в задачу разработки адаптивных обучающих программ, которые были бы в состоянии приспосабливаться к каждому обучаемому, руководствуясь историей его предшествующего обучения. Следовательно, речь должна идти о регистрации тех результатов обучения, которые дадут информацию о возможных путях индивидуализации этого процесса.

Не лучше обстоит дело с решением вопроса о том, как измерять. "При планировании методов обработки учебных показателей нескольких тысяч стэнфордских школьников оказалось, что труднее всего решить, какие показатели из общей массы имеет смысл регистрировать постоянно. Когда за учебными стендами занимаются 5000 школьников, никакого труда не представляет с помощью вычислительной машины получать ежедневные отчеты о их работе толщиной в 1000 страниц, но регулярно перерабатывать такое количество информации выше человеческих сил. 1000 страниц необходимо сократить до 25–30. ...Вся трудность заключается в том, что не имеется убедительных теоретических построений, которые могли бы указать, как надо проводить такое сокращение. В настоящее время все решения, которые мы принимаем в подобных случаях, базируются на педагогической интуиции и традиционном анализе данных, принятом в экспериментальной психологии. Ни то, ни другое не может служить надежным показателем в учебной работе" [13, с. 181 – 182].

В целом приходится признать, что задача разработки дидактической квалиметрии, несмотря на ее практическую актуальность, еще только ставится. "В педагогике предстоит разрабо-

тать и обосновать дидактически независимую систему основных и производных единиц измерения. Вместе с тем важно определить, какие характеристики учебных объектов являются исходными и подлежат первичному измерению" [7, с.75].

Тем не менее разработчикам методического обеспечения АОС, занимающимся адаптивными обучающими программами, приходится принимать конкретные практические решения по вопросам дидактической квалиметрии. Без этого функционирование адаптивной обучающей программы в АОС просто невозможно. Результаты, полученные нами [2], позволяют критически подойти к довольно прочно устоявшимся в дидактике догмам. Речь пойдет о том, что в поисках ответа на вопросы дидактических измерений исследователи, как правило, обращаются к теории психологических измерений, получившей свой метрологический статус и находящей себе все более широкое применение [10, 11, 16]. Целесообразно остановиться на основных положениях этой теории и попытаться установить границы ее применимости к дидактическим измерениям при машинном обучении.

"В самом общем смысле слово "измерение" обозначает операцию, посредством которой числа (или по крайней мере порядковые величины) приписываются вещам" [2, с.195]. Несмотря на внешнюю простоту и инструктивную ясность этого определения, введение его в практику психологических измерений сопряжено с рядом существенных трудностей. Все они проистекают из того, что "вещь" которой следует приписать числа, - это личность. Если в естественных науках измерение, как правило, осуществляется фиксацией всех переменных, кроме одной - независимой, представляющей собой измеряемую величину, то в психологии любое явление зависит обычно от многих факторов. Отсюда возникает острый вопрос применения методов многомерного анализа психологических явлений [4] и отсюда же - вопрос перехода от качественных шкал к адекватному измерению множества свойств в их количественном выражении.

С методологической точки зрения для любого измерения необходимо выбрать единицу анализа. В психологических исследованиях измерению подвергается предметная деятельность испытуемого или какая-то составная часть ее. Еще чаще о ее протекании судят по конечному продукту. Сложившиеся в психологии представ-

ления о предметной деятельности дают следующую ее схему: деятельность - действие - операция [5]. При этом деятельность не является аддитивным процессом, состоящим из простой суммы действий. Если при измерении из предметной деятельности исключить осуществляющие ее действия, то от деятельности ничего не останется. По-видимому, на такой методологической основе строит измерение качества обучения В.П.Беспалько [3, с.199], который практически ставит знак равенства между понятиями усвоения деятельности и действия. Исходя из психологической теории деятельности, измерить усвоение деятельности невозможно вообще, ибо само понятие деятельности в психологии выступает как универсальная предельная абстракция [9].

Если измерять операцию, т.е. способ осуществления действия, то как найти некоторую алгебраическую функцию, количественно отражающую эту качественную сторону деятельности? Если же измерять действие, т.е. процесс, подчиненный сознательной цели, - то сразу же возникает необходимость выйти за пределы трехчленной структуры предметной деятельности в сферу целеполагания. Иначе, измеряя лишь чисто внешнюю, предметную сторону действия, исследователь рискует упустить ее психологический субстрат. Таким образом встает вопрос: как измерить мотивационную сферу деятельности, представленную А.Н.Леонтьевым в виде второго ряда понятий: мотив - цель - условие? [5, с. 102 - 105]. Мотив является фактором, определяющим выбранный путь к цели и извлекающим из памяти весь прошлый опыт с коррекцией на существующую обстановку [1, с. 9]. В одной ситуации обучаемый сочтет целесообразным обратиться к учебнику или обучаемой системе, в другой - получить консультацию у преподавателя. В обоих случаях разные пути приведут к одной и той же цели, но как измерить мотив выбранного решения? Непосредственному измерению доступна лишь экстериоризованная деятельность, в основе которой лежат вербальные или двигательные компоненты. Соотнесения понятий мотива и деятельности, а также понятий цели и действия явно недостаточно для установления соответствия между свойствами этих явлений и свойствами чисел. Это усугубляется еще и тем, что понятие мотива определяется неоднозначно; дополнительно вводится еще одно определение мотива-цели как осознанной генерализованной цели деятельности.

Как видно, встающие здесь вопросы прямо связаны с содержанием психологической теории. Ответ на них порой зависит от той концепции, которой придерживается исследователь. С математической же точки зрения от них можно абстрагироваться, и тогда психологические измерения сведутся к выбору вида шкал, т.е. определению способа перехода от одной числовой системы к адекватной системе с теми же отношениями. Другими словами, требуется отыскать возможность психологического и числового прочтения одних и тех же формальных отношений в рассматриваемых явлениях. Из аксиом сравнения величин [16] вытекает существование четырех видов шкал [14, 15]: шкал наименований, шкал порядка, шкал интервалов, шкал отношений.

Шкалы наименований устанавливают отношения равномерности, которые распределяют экспериментальные данные по определенному числу классов. Каждый результат относится только к одному классу. Примером может служить классификация профессий по признаку взаимодействия человека с окружающей действительностью - профессии: человек - человек, человек - природа, человек - знак, человек - образ, человек - техника. К этим же шкалам может быть отнесено выделение элементов деятельности человека: сенсорный компонент деятельности, речевой, моторный и т.д. Числам, обозначающим эти классы, достаточно обладать только одним свойством - быть отличными друг от друга.

Шкалы порядка устанавливают между явлениями отношения равенства и отношения ранговой последовательности. Числа, приписываемые шкалам порядка, должны иметь свойства монотонного преобразования, т.е. свойства остаются неизменными при замене этих чисел упорядоченным рядом чисел. Кроме того, за этими числами должно сохраняться и свойство отличительности их друг от друга, необходимое при составлении шкалы наименований. В качестве примера шкалы порядка можно привести ранжирование одних и тех же качеств при оценке и самооценке личности; субъективное упорядочивание звуковых, кинестезических, зрительных или тактильных стимулов по их возрастающей (убывающей) интенсивности. В частности, подсчет коэффициента корреляции рангов по Спирмену основан на том допущении, что числа, образующие ранги, обладают свойством отличительности признаков и

свойствами монотонного преобразования.

Шкалы порядка могут также применяться для выявления результатов обучения. Расположив обучаемых последовательно по признаку скорости выработки навыков, можно судить об их большей или меньшей способности к обучению. По характеру кривой распределения статистических результатов представляется возможным выявить такую меру центральной тенденции, как моду, характеризующую максимальную вероятность появления ожидаемых результатов. В психологическом смысле это означает, что в конкретной изучаемой деятельности человека уже сформировался навык (действие или операция). В педагогическом аспекте это же может характеризовать успешность обучения.

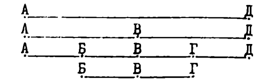
В проведенных нами исследованиях [6] были поставлены эксперименты с целью установления характера статистического распределения при формировании сложного навыка. Испытуемые проходили тренировку в автоматизированной обучающей системе. Предполагалось, что по мере выработки сложного навыка на ограниченном временном отрезке медиана и мода как меры центральной тенденции будут сближаться. При математической обработке полученные 6300 замеров были ранжированы в определенной последовательности и проанализированы. Гипотеза о сближении моды и медианы по мере формирования навыка полностью подтвердилась. Таким образом, применение шкалы порядка в оценке результатов психолого-педагогического эксперимента позволило ограничиться достаточно простой измерительной техникой и оказалось достаточно эффективной мерой для определения обучаемости в целом.

В экспериментах можно не вводить ограничение времени, а обучаемым позволить осваивать данную деятельность до полного овладения ею. Тогда, расположив испытуемых по признаку максимальных затрат времени, можно также определить их обучаемость. На этом основаны известные шкалы умственных способностей, предложенные Бинэ, которые относятся к рассматриваемому типу.

Шкалы интервалов строятся на основе одинаковой разности между двумя парами каких-либо фактов, т.е. двум экспериментально равным различиям приписываются два равных числовых значения. Одна из экспериментальных операций, определяющих равновеликую разность между двумя парами фактов,

схематично может быть представлена так.

В некотором измеряемом явлении известны два факта или точки: А и Д (рисунок). Вначале этот интервал делится точкой В на два равных интервала АВ и ВД. Затем точно так же определяются точки Б и Г, соответственно делящие на равные части отрезки АВ и ВД. Последняя контрольная операция – установить совпадение середины отрезка БГ с точкой В. Сейчас остается только предполагать, что так была построена стоградусная температурная шкала Цельсия, где за нуль взята точка замерзания воды, а за 100° – точка ее кипения при нормальном давлении. Применение этой физической шкалы в психофизиологических измерениях затруднительно, ибо геометрически равные интервалы не соответствуют субъективным интервалам в ощущениях индивида. Даже в незначительном диапазоне изменение температуры на один градус имеет совершенно разную значимость для организма в зависимости от того, произошло оно от 36,6 или от 41°С.



Принцип построения шкалы интервалов

Свойства чисел, приписываемых шкалам интервалов, сохраняются неизменными после линейного преобразования вида $y=kx+b$, а измерение предполагает линейность измеряемого явления. В таких шкалах выбирается произвольно начало (параметр b) и единица (параметр k), как например, в температурных шкалах Фаренгейта, Реомюра и Цельсия. В оценочных суждениях или субъективных оценках индивида значимость изменяющегося свойства или явления нелинейна, поэтому требуется более сложная математическая конструкция.

Из области психофизики убедительной иллюстрацией к сказанному будет закон Вебера – Фехнера (основной психофизический закон), устанавливающий логарифмическую зависимость между равными интервалами изменения субъективных ощущений и физическими стимулами, вызывающими эти ощущения. В педагогике нелинейность применяемых шкал также имеет место. Каждому экзаменатору из собственной практики известно, сколь трудно бывает оценить знания обучаемого по существующей четырех-

балльной шкале, если качество этих знаний не лежит в областях, пограничных с "неудовлетворительно" и "отлично". При этом априори считаются равновеликими интервалы между двумя любыми соседствующими баллами, хотя содержание качества знаний умений или навыков, подлежащих оценке конкретным баллом в высшей школе, нигде не оговорено.

Не меньшие трудности возникают в спорте, где оценка мастерства спортсменов проводится по шкале интервалов с одинаковыми числовыми различиями между двумя фактами. К этим видам спорта можно отнести фигурное катание, художественную гимнастику, прыжки в воду, акробатику и т.д. Перечень подобных примеров можно было бы продолжить, заимствовав их из сферы искусства.

Во всех перечисленных случаях для большей объективности применяется ставший традиционным метод экспертных оценок с последующей статистической обработкой. Отыскивается средняя арифметическая и отклонения от нее, позволяющие устранить слишком противоречивые суждения об одних и тех же фактах.

Ш к а л ы о т н о ш е н и й строятся на основе равенства отношений между двумя парами фактов, что можно представить в виде: факт А некоторого явления количественно относится к факту Б другого явления так же, как факт В третьего явления относится к факту Г четвертого явления. Если это экспериментально подтверждается, то исследователь вправе приписать изучаемым явлениям числа, отношения между которыми равны. Свойства этих чисел должны сохраняться в результате преобразования вида $y = kx$, и применительно к ним все арифметические операции имеют смысл. Например, количество прочитанных страниц научной литературы x актуализируют уже имеющиеся знания y в k раз больше, чем информация, заложенная в числе страниц x. Если сравнить преобразования этого вида с преобразованиями, которым подчиняются шкалы интервалов, то видно, что они отличаются параметром b. Исчезновение параметра b означает не произвольность, а жесткую фиксированность начала отсчета в шкалах отношений.

Шкалирование может применяться всюду, где изучается некоторое свойство качественной природы при условии упорядочения его структуры. Тогда задача измерения сводится к отобра-

жению этого свойства в числовом множестве. В зависимости от свойств полученных исходных множеств, могут применяться и производные шкалы, цель которых – численно описать факты, вскрытые первичной статистической обработкой. Сюда можно отнести методы вычисления корреляций и факторный анализ. При любом количественном описании психических явлений всегда сохраняется требование надежности выполняемых измерений. Смысл надежности измерений имеет два аспекта.

Во-первых, измерение должно быть применимо к объекту. Свойства измеряемого объекта или явления не должны быть зависимы от инструмента и процедуры измерения. Инструменты измерения, не обладающие постоянством собственных свойств по отношению к различным испытуемым, в исследованиях не применимы. Так и психические явления обследуемого индивида, в том числе и обучаемого, не должны изменять характер своего протекания под влиянием применяемой аппаратуры или тестов. При этих условиях достоверные результаты дает многократное измерение. Здесь вполне допустима систематическая ошибка, инвариантная при повторных замерах. Математическая интерпретация подобных ошибок в шкалах с произвольным началом (в шкалах интервалов) будет представлять собой смещение исходной отметки на величину $в$.

Другой аспект касается воспроизводимости существующих зависимостей между стимулом и реакцией на этот стимул, т.е. обнаруженную зависимость требуется подтвердить многократно. Численная обработка должна основываться на многочисленных измерениях. Количественная проверка надежности осуществляется методами, описание которых можно найти в любом пособии по математической статистике. Один из самых простых критериев надежности измерений известен как формула Кадера – Ричардсона [17, с. 318].

Таким образом, теория шкал измерения касается природы исследуемого объекта лишь в той мере, в какой последняя влияет на выбор шкалы. В этом смысле теория измерений универсальна, она приложима к объектам исследования любой природы. Теория же психологических измерений подобной универсальностью не обладает. Это учение о выборе шкал измерения при исследовании психологических явлений. Решение вопроса о том, ч т о из-

мерять, остается прерогативой содержательных представлений исследователя об исследуемом объекте. Только на их основе можно решать вопрос,"что измерять", как и вообще вопрос применимости математических методов в конкретной науке [4, 8, 15]. Лишь после этого наступает очередь выбора шкалы измерения.

Уяснение этого обстоятельства позволяет вновь обратиться к проблеме дидактических измерений при машинном обучении. В чем слабость психологического подхода к измерениям в дидактике? В психологии измерению подвергается деятельность, об успешности протекания которой судят по ее конечному или промежуточному продукту. Применительно к обучению одни и те же измерения устанавливают факты протекания процесса и получения результата. Налицо смешение понятий, что, вообще говоря, следует признать источником многих недоразумений и в психологии и в дидактике. "Неумение провести различие между процессом обучения и критериями обучения, - отмечает Х.Хелсон, - ведет к большой путанице. Часто способ обучения трактуется так, будто условия, метод или путь, с помощью которых обучение имеет место, образуют обучение в смысле того, что́ достигается" [18, с. 392].

Нам представляется, что без такого разграничения проблема дидактических измерений не может быть поставлена корректно; специфика дидактических явлений в сравнении с психологическими заключается, в частности, в несводимости результатов педагогического процесса к чисто психической интериоризованной деятельности обучаемого. В конечном счете обучение имеет своей целью социальную подготовку обучаемых к деятельности в сфере общественного производства. Следовательно, при изучении дидактических явлений во внимание должны приниматься не только свойства личности обучаемого, но и объективно существующие нормы производственной деятельности в системе разделения и кооперации общественного труда.

Такой подход означает прежде всего изменение содержательного представления об обучении. Обучение начинает рассматриваться как процесс адаптации к объективно существующим нормам производственной деятельности, а результатом этого процесса признается совпадение деятельности обучаемого с внешними для него нормами практической деятельности людей в сфере общест-

венного производства. Вместе с этим изменяется постановка задачи дидактических измерений, ибо в поле зрения исследователя оказывается уже социально обусловленное явление. Разграничение задач психологических и дидактических измерений совершенно необходимо при адаптивном обучении с помощью АОС. Этого разграничения нельзя, по-видимому, избежать и при решении проблемы дидактических измерений в целом.

Литература

1. Анохин П. К. Проблема принятия решения в психологии и физиологии. - В кн.: Проблемы принятия решения. М., Наука, 1976.

2. Белякин А. М., Занько С. Ф., Яхонтов В. Н. Об одном алгоритме взаимодействия обучаемого с автоматизированной обучающей системой в адаптивном режиме обучения. - В сб.: Исследования по техническим обучающим системам. Казань, 1978 (Казанский авиационный институт).

3. Беспалько В. П. Основы теории педагогических систем. Проблемы и методы психолого-педагогического обеспечения технических обучающих систем. Воронеж, Изд-во Воронежского университета, 1977.

4. Битинас Б. П. Методы многомерного анализа в педагогическом исследовании. - В кн.: Методы педагогических исследований в высшей школе. Уфа, 1976 (Уфимский авиационный институт).

5. Леонтьев А. Н. Деятельность. Сознание. Личность. М., Политиздат, 1975.

6. Михайлик Н. Ф., Перфильев А. К., Юсупов И. М. К статистической оценке формирования сложного навыка обучаемых автоматизированными обучающими системами. - В сб.: Исследования по техническим обучающим системам. Казань, 1978 (Казанский авиационный институт).

7. Мизинцев В. П. Пути исследования проблемы количественной оценки эффективности учебного процесса. Советская педагогика, 1979, № 8.

8. Моисеев Н. Н. Математика ставит эксперимент. М., Наука, 1979.

9. Огурцов А. П., Юдин Э. Г. Деятельность. БСЭ, т. 8. М., Советская энциклопедия, 1970.

10. Психологические измерения. М., Мир, 1976.

11. Пфанцагль И. Теория измерений. М., Мир, 1967.

12. Рёшлен М. Измерения в психологии. - В кн.: Экспериментальная психология, вып. 1 - 2. Под редакцией П. Фресса и Ж. Пиаже. М., Прогресс, 1966.

13. Саппэс П. Образование и вычислительные машины.-

В кн.: Информация. М., Мир, 1968.

14. Стивенс С. С. Экспериментальная психологи
т. I, 2. М., ИЛ, 1960.

15. Фридман Л. М. Использование моделирования
измерения в психолого-педагогических исследованиях. – В кн.
Методы педагогических исследований в высшей школе. Уфа, 1976
(Уфимский авиационный институт).

16. Холл А. Опыт методологии для системотехники. М.
Сов. радио, 1975.

17. Ebel R. Measuring Educatinal achievement. Nev
Jersey, 1965.

18. Helson H. Adaption-Level Theory. New York, 196

Поступила в редколлеги
19 июня 1979г.